Appl. No. 09/683,264
Amdt. dated November 24, 2004
Amendment under 37 CFR 1.116 Expedited Procedure
Examining Group

PATENT

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1.   (Withdrawn) A computer implemented method for characterizing a plurality of biological sequences comprising:

obtaining a plurality of models, wherein each of the models represents a classification of biological sequences with structural or functional similarity; determining fitness of the biological sequences to the models; and automatically classifying the sequences according to the distances to the models.

2.   (Withdrawn) The method of Claim 1 wherein the plurality of biological sequences have at least 50 sequences.

3.   (Withdrawn) The method of Claim 2 wherein the plurality of biological sequences have at least 100 sequences.

4.   (Withdrawn) The method of Claim 3 wherein the plurality of biological sequences have at least 100 sequences.

5.   (Withdrawn) The method of Claim 3 wherein the models are Hidden markov models.

6.   (Withdrawn) The method of Claim 5 wherein the classification is a family and each model represents a family.

7.   (Withdrawn) The method of Claim 6 wherein the sequences are protein sequences.

8.   (Withdrawn) The method of Claim 7 wherein the distances are E values.

Appl. No. 09/683,264
Amdt. dated November 24, 2004
Amendment under 37 CFR 1.116 Expedited Procedure
Examining Group

PATENT

9.   (Withdrawn) The method of Claim 8 wherein the step of automatically determining comprises determining a step of determining a threshold for each of the models.

10.   (Withdrawn) The method of Claim 9 wherein the step of determining a threshold comprises performing a curve analysis.

11.   (Withdrawn) The method of Claim 10 wherein the step of performing a curve analysis comprises determining a point where the e-value curve drops abruptly or flattens.

12.-19.   (Canceled)

20.   (Withdrawn) A system for gene annotation comprising a processor; and a memory coupled with the processor, the memory storing a plurality of machine instructions that cause the processor to perform logical steps comprising obtaining a plurality of models, wherein each of the models represents a classification of biological sequences with structural or functional similarity; determining fitness of the biological sequences to the models; and automatically classifying the sequences according to the distances to the models.

21.   (Withdrawn) The system of Claim 20 wherein the plurality of biological sequences have at least 50 sequences.

22.   (Withdrawn) The system of Claim 21 wherein the plurality of biological sequences have at least 100 sequences.

23.   (Withdrawn) The system of Claim 22 wherein the plurality of biological sequences have at least 100 sequences.

24.   (Withdrawn) The system of Claim 23 wherein the models are Hidden markov models.

25.   (Withdrawn) The system of Claim 24 wherein the classification is a family and each model represents a family.

26.   (Withdrawn) The system of Claim 25 wherein the sequences are protein sequences.

27.   (Withdrawn) The system of Claim 26 wherein the distances are E-values.

28.   (Withdrawn) The system of Claim 27 wherein the step of automatically determining comprises determining a step of determining a threshold for each of the models.

29.   (Withdrawn) The system of Claim 28 wherein the step of determining a threshold comprises performing a curve analysis.

30.   (Withdrawn) The system of Claim 29 wherein the step of performing a curve analysis comprises determining a point where the e value curve drops abruptly or flattens.

31.-38.   (Canceled)

39.   (Withdrawn) A computer software product of the invention comprising a computer readable medium having computer-executable instructions for performing the method comprising:

obtaining a plurality of models, wherein each of the models represents a classification of biological sequences with structural or functional similarity;

determining fitness of the biological sequences to the models; and

automatically classifying the sequences according to the distances to the models.

40.   (Withdrawn) The product of Claim 39 wherein the plurality of biological sequences have at least 50 sequences.

41.   (Withdrawn) The product of Claim 40 wherein the plurality of biological sequences have at least 100 sequences.

42.   (Withdrawn) The product of Claim 41 wherein the plurality of biological sequences have at least 100 sequences.

43.   (Withdrawn) The product of Claim 42 wherein the models are Hidden markov models.

44.   (Withdrawn) The product of Claim 43 wherein the classification is a family and each model represents a family.

45.   (Withdrawn) The product of Claim 44 wherein the sequences are protein sequences.

46.   (Withdrawn) The product of Claim 45 wherein the distances are E-values.

47.   (Withdrawn) The product of Claim 46 wherein the step of automatically determining comprises determining a step of determining a threshold for each of the models.

48.   (Withdrawn) The product of Claim 47 wherein the step of determining a threshold comprises performing a curve analysis.

49.   (Withdrawn) The product of Claim 48 wherein the step of performing a curve analysis comprises determining a point where the e-value curve drops abruptly or flattens.

50.-81.   (Canceled)

82.   (New) A computer implemented method for gene characterization comprising:

generating a plurality of models using structural relationships of known proteins;

inputting a plurality of protein sequences;

determining a plurality of scores by comparing the plurality of protein sequences with the plurality of models;

automatically selecting a plurality of hits based on at least the plurality of scores and a plurality of criteria; and

assigning the plurality of protein sequences to the plurality of models based on at least the plurality of selected hits.

Appl. No. 09/683,264
Amdt. dated November 24, 2004
Amendment under 37 CFR 1.116 Expedited Procedure
Examining Group

PATENT

83.   (New) The method of claim 82 wherein the plurality of models comprises hidden markov models.

84.   (New) The method of claim 82 wherein the plurality of protein sequences comprises 50 protein sequences.

85.   (New) The method of claim 84 wherein the plurality of protein sequences comprises 150 protein sequences.

86.   (New) The method of claim 85 wherein the plurality of protein sequences comprises 500 protein sequences.

87.   (New) The method of claim 82 wherein the automatically selecting a plurality of hits comprises determining a threshold for each of the plurality of models.

88.   (New) The method of claim 87 wherein the determining a threshold comprises performing a curve analysis.

89.   (New) The method of claim 88 wherein the performing a curve analysis comprises analyzing a plurality of slopes for a plurality of curves.

90.   (New) A system for gene characterization comprising a processor; and a memory coupled with the processor, the memory storing a plurality of machine instructions that cause the processor to perform logical processes comprising:

    generating a plurality of models using structural relationships of known proteins;

    inputting a plurality of protein sequences;

    determining a plurality of scores by comparing the plurality of protein sequences with the plurality of models;

    automatically selecting a plurality of hits based on at least the plurality of scores and a plurality of criteria; and

Appl. No. 09/683,264
Amdt. dated November 24, 2004
Amendment under 37 CFR 1.116 Expedited Procedure
Examining Group

PATENT

assigning the plurality of protein sequences to the plurality of models based on at least the plurality of selected hits.

91.   (New) The system of claim 90 wherein the plurality of models comprises hidden markov models.

92.   (New) The system of claim 90 wherein the plurality of protein sequences comprises 50 protein sequences.

93.   (New) The system of Claim 92 wherein the plurality of protein sequences comprises 150 protein sequences.

94.   (New) The system of claim 93 wherein the plurality of protein sequences comprises 500 protein sequences.

95.   (New) The system of claim 90 wherein the automatically selecting a plurality of hits comprises determining a threshold for each of the plurality of models.

96.   (New) The system of Claim 95 wherein the determining a threshold comprises performing a curve analysis.

97.   (New) The system of Claim 96 wherein the performing a curve analysis comprises analyzing a plurality of slopes for a plurality of curves.

98.   (New) A computer software product for gene characterization comprising a computer readable medium having computer executable instructions for performing the method comprising:

generating a plurality of models using structural relationships of known proteins;

inputting a plurality of protein sequences;

determining a plurality of scores by comparing the plurality of protein sequences with the plurality of models;

Appl. No. 09/683,264
Amdt. dated November 24, 2004
Amendment under 37 CFR 1.116 Expedited Procedure
Examining Group

PATENT

automatically selecting a plurality of hits based on at least the plurality of scores and a plurality of criteria; and

assigning the plurality of protein sequences to the plurality of models based on at least the plurality of selected hits.

99. (New) The product of claim 98 wherein the plurality of models comprises hidden markov models.

100. (New) The product of claim 98 wherein the plurality of protein sequences comprises 50 protein sequences.

101. (New) The product of claim 100 wherein the plurality of protein sequences comprises 150 protein sequences.

102. (New) The product of claim 101 wherein the plurality of protein sequences comprises 500 protein sequences.

103. (New) The product of claim 98 wherein the automatically selecting a plurality of hits comprises determining a threshold for each of the plurality of models.

104. (New) The product of claim 103 wherein the determining a threshold comprises performing a curve analysis.

105. (New) The product of claim 104 wherein the performing a curve analysis comprises analyzing a plurality of slopes for a plurality of curves.